NIST MSPepSearch

(for batch interpretation of peptide MS/MS spectra)

Version 2009 (v. 0.9) Build 05/26/09
for use with Microsoft© Windows™
by

Mass Spectrometry Data Center
Director, Stephen E. Stein
National Institute of Standards and Technology

Background and Intended Use

This program has been written with proteomics users in mind. This program uses a special version of the NIST MS Search 2.0 engine (Nistdl32a.dll) to quickly identify the best matches to a set of unknown spectra in a peptide mass spectral library (in NIST binary format). Search logistics and scoring have been appropriately modified for use with large peptide fragmentation mass spectral libraries. MSPepSearch currently reads input spectra in MGF format (http://www.matrixscience.com/help/data_file_help.html) and outputs a match list in tab-delimited text. The output can be easily viewed in a spreadsheet application or parsed with scripts. Since the program runs in "batch" mode, it is suitable for use in data analysis pipelines. Search libraries are compatible with NIST MS Search 2.0, enabling MSPepSearch matches to be viewed with the MS Search 2.0 interface. A demo of this program can be downloaded here (http://chemdata.nist.gov/mass-spc/Srch_v1.7/NISTDEMO_08.EXE). Note: The demo software does not contain the 2008 NIST/EPA/NIH Mass Spectral Library (http://www.nist.gov/srd/nist1a.htm) for interpretation of EI mass spectra.

This program is intended for evaluation and use by developers and researchers who are interested in identification of peptides (i.e., mainly those generated by digestion with trypsin) by spectral library searching. To download libraries or for more information, see http://peptide.nist.gov.

Support contact: Paul Rudnick (*paul.rudnick@nist.gov*)

Programmatic contact: Steve Stein (*steve.stein@nist.gov*)

Programming Credits: Dmitrii Tchekhovskoi, Steve Stein and Yuri Mirokhin

Quick Installation Instructions

1. *Extract the MSPepSearch.zip archive in a working directory on your hard drive (e.g., C:\projects\MSPepSearch).*
2. *Download one or more peptide libraries in 'NIST' format from* http://peptide.nist.gov and save them to your working directory.
3. *Open a DOS command windows by Clicking Start->Run and typing 'cmd'.*

4. *Run the program by pointing to an MGF file containing unknown spectra and 1-4 spectral libraries. Here is an example command for searching data from an ion trap mass spectrometer such as a Thermo Scientific© LTQ™:*

```
C:\projects\MSPepSearch\MSPepSearch.exe fiPv /OutSpecNum /HiPri
/LibInMem /MinMF 450 /HITS 3 /Z 1.6 /M 0.8 /LIB
"C:\projects\MSPepSearch\yeast_consensus_final_true_lib" /LIB
"C:\projects\MSPepSearch\human_consensus_final_true_lib" /INP
"C:\projects\MSPepSearch\test-spectra.RAW.MGF" /OUTTAB
"C:\projects\MSPepSearch\test-spectra.RAW.MGF.TAB"
```

The /Z parameter sets the precursor tolerance and the /M parameter sets the fragment tolerance (both are in *m/z*). /MinMF and /HITS limit the output file to Scores >=450 (a reasonable similarity threshold) and to the top 3 matches per spectrum, respectively. Libraries are specified by the full directory name. The input file is expected in MGF format; its name must have .MGF extension, as in 'test-spectra.RAW.MGF'.

In the output file, 'Score' is the most discriminant score and should be used to assess false discovery rates.

Restrictions and Disclaimers

MSPepSearch binaries and dlls may be redistributed without restriction. Peptide mass spectral libraries are freely available, however the libraries may not be redistributed without permission.

The National Institute of Standards and Technology (NIST) uses its best efforts to deliver high-quality reference data and software products. However, NIST makes no warranties to that effect; and NIST shall not be liable for any damage that may result from the use of the above mentioned products.

Certain commercial equipment, instruments, or materials are identified in this document. Such identification does not imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the products identified are necessarily the best available for the purpose.

NIST
National Institute of
Standards and Technology
U.S. Department of Commerce